

Eur. J. Clin. Chem. Clin. Biochem.  
Vol. 31, 1993, pp. 165–172

© 1993 Walter de Gruyter & Co.  
Berlin · New York

## Evaluation of a Fully Mechanised Immunoassay — Enzymun-Test System ES 300 — and Comparison with In-House Methods for 8 Analytes

By Marion Bauernfeind<sup>1</sup> and W. G. Wood<sup>2</sup>

<sup>1</sup> Immunodiagnostic Evaluation Dept., Boehringer Mannheim, Mannheim, Germany

<sup>2</sup> Institut für Klinische Laboratoriumsdiagnostik, Klinikum der Hansestadt Stralsund, Stralsund, Germany

(Received June 22/October 5, 1992)

**Summary:** The following analytes were determined with the automated enzyme immunoassay unit, Enzymun-Test System ES 300, and by routine in-house tests, and the results were compared: thyrotropin, free thyroxine, cortisol, immunoglobulin E, digoxin, digitoxin, insulin and carcinoembryonic antigen. The methods used for comparison included two radioimmunoassays, one with enzyme labelling, two with fluorescence enhanced enzyme reactions and 3 luminescence immunoassays. In most cases, the precision of the ES 300 lay between 2 and 5% in all assays in the concentration ranges of clinical interest. The inter-assay variation was almost identical to the intra-assay precision, which reflects the constancy of reagent quality and machine performance.

The correlation between in-house and Enzymun-Test was excellent, even though the slope of the regression line was sometimes far from unity, due to the calibration materials used.

ES 300 can be used as a stand alone unit and connected to a host computer. As a “walk-away” machine, it is suitable for laboratories with medium length series. It is a multi-batch analyser capable of taking 12 analytes and a maximal series length of 136 tubes excluding standards and controls.

### Introduction

Enzymun-Test System ES 300 is a fully automated enzyme immunoassay analysis system consisting of a pipettor, incubator, photometer and electronic data processing (EDP)-unit. Interfacing to a central or host computer is possible as is a manual input via a conventional PC-keyboard. Enzymun-Test System ES 300 is a bench-top analyser, related to Enzymun-Test System ES 600 from the same manufacturer (1), and it is a stand-alone unit. Enzymun-Test System ES 300 is a patient orientated multibatch analyser designed as a workhorse for middle sized laboratories and it can incorporate from 1 to 12 analytes and up to 150 samples in one run.

### Materials and Methods

Table 1 shows the list of analytes tested, together with the methods used for comparison.

Tab. 1. List of analytes tested together with methods used for comparison

Analyte	Method of comparison*)	Producer of Kit
CEA	IEMA	Abbott — EIA Monoclonal
Cortisol	RIA	Diagnostic Products-Coat-a-Count
Thyrotropin	ILMA	Behringwerke — BeriLux
Free T <sub>4</sub>	LIA	Ciba-Corning — Magic Lite
Digoxin	FEIA	Baxter — Stratus
Digitoxin	FEIA	Baxter — Stratus
IgE	ILMA	In-house assay
Insulin	RIA	Medgenix

\*) Key: RIA — Radioimmunoassay  
LIA — Luminescence Immunoassay  
IEMA — Immunoenzymometric Assay  
ILMA — Immunoluminometric Assay  
FEIA — Fluorescence Enhanced Enzyme Immunoassay

Tab. 2. Reference — or therapeutic ranges given by the kit producer<sup>1</sup> or established in the laboratory<sup>2</sup>

Analyte	Reference/Therapeutic range		Method used for comparison	
	Enzymun-Test System ES 300			
CEA	<4.9 (<2.65)	$\mu\text{g/l}^a$ $\mu\text{g/l}^b$	<2.5	$\mu\text{g/l}^b$
Cortisol	70 — 250	$\mu\text{g/l}^c$	40 — 200	$\mu\text{g/l}^c$
Thyrotropin	0.25 — 3.10	mU/l <sup>a</sup>	0.20 — 3.50	mU/l <sup>b</sup>
Free T <sub>4</sub>	8 — 20	ng/l <sup>a</sup>	8 — 20	ng/l <sup>b</sup>
Digoxin	0.5 — 2.0	$\mu\text{g/l}^a$	0.8 — 2.2	$\mu\text{g/l}^b$
Digitoxin	9 — 30	$\mu\text{g/l}^a$	8 — 25	$\mu\text{g/l}^b$
IgE	<100	kU/l <sup>d</sup>	<20	kU/l <sup>e</sup> (negative)
			20 — 100	kU/l <sup>e</sup> (borderline)
Insulin	3 — 17	mU/l <sup>a</sup>	5 — 12	mU/l <sup>b</sup>

<sup>a</sup> CEA values irrespective of smokers/non-smokers<sup>b</sup> CEA values established on healthy non-smokers<sup>c</sup> Cortisol values for 7.00 am.<sup>d</sup> IgE values above 100 kU/l indicate an allergy<sup>e</sup> IgE values under 20 kU/l for normal individuals; values between 20–100 kU/l do not exclude an allergy.

Table 2 shows the reference or therapeutic ranges for the methods used as declared by the kit producers.

The in-house method for IgE used antibodies from DAKO — Hamburg, D and Cappel — Turnhout, B. The label used was 9-(4-aminobutyl-N-ethyl) benzo(f)phthalazine -1,4 (2H, 3H) dione (ABEN) coupled to the DAKO antibody.

The other kits were used according to the manufacturers' instructions.

The Baxter-Stratus was a closed system; the Abbott IEMA was read off on an Abbott Quantum II manual EIA-reader with microprocessor and printer.

For the radioisotopically labelled kits, counting was performed in an NE-1600 16-well manual gamma spectrometer (Thorn-EMI Sighthill, Edinburgh UK).

For the kits with luminescent labels, measurements were performed in a Berthold LB-952 16 T 250 sample chain luminometer with microprocessor and printer (Berthold-EG & G, Wildbad, D).

The chromogen used in the Enzymun-Test kit was 2,2'-azino-di [3-ethylbenzthiazoline 6-sulphonic acid], (ABTS).

Non-parametric statistics were used throughout, regression lines being constructed using the *Passing-Bablok* main-component analysis method (2).

### Concepts and Targets of the Evaluation

#### Precision and accuracy

For all methods, inter- and intra-assay coefficients of variation were determined on pool sera. Three human serum pools of different concentration and two control sera from Boehringer Mannheim were used for estimation of inter- and intra-assay CV for ES 300. Accuracy could not be tested, due to the absence of reference or definitive methods. The comparison between ES 300 and the in-house assay may be taken as a relative estimate of the accuracy by using the slope of the regression line and intercept value.

#### Lower detection limit (tab. 3) and intra-assay drift

The potential lower detection limit of the in-house assays was determined from the 99% confidence limit from the zero stan-

Tab. 3. Lower detection/decision limits for quantitative documentation in the Enzymun-Test System ES-300 and in-house methods

Analyte	Enzymun-Test System ES 300	In-house method
CEA	0.5 $\mu\text{g/l}$	1 $\mu\text{g/l}$
Cortisol	2 $\mu\text{g/l}$	5 $\mu\text{g/l}$
Thyrotropin	0.07 mU/l	0.005 mU/l
Digoxin	0.3 $\mu\text{g/l}$	0.2 $\mu\text{g/l}$
Digitoxin	2 $\mu\text{g/l}$	2 $\mu\text{g/l}$
IgE	1.5 kU/l	10 kU/l
Insulin	3 mU/l	5 mU/l
Free T <sub>4</sub>	2 ng/l	1.5 ng/l

Values lower than those in the above table were given as < or "Less than".

dard (+3 s for immunometric assays, -3 s for competitive assays), derived from duplicate results in 10 consecutive assays.

The lower detection limits used in the ES 300 evaluation were not calculated during evaluation, but were taken from internal and external evaluation studies. The intra-assay drift was calculated from 25 identical samples placed at 4 tube intervals during a run.

### Clinical relevance of results

In addition to the classification of results obtained from both methods, sera giving widely divergent results in both methods were investigated further for icterus, haemolysis and lipaemia, taking into account the accompanying clinical diagnosis and, where known, possible drug interferences.

### Results and Discussion

Regression analysis — linear regression according to *Passing-Bablok* (2)

*Carcinoembryonic antigen* — CEA (fig. 1)

Patient samples (n = 140) were measured simultaneously with both tests. Twenty runs were performed,

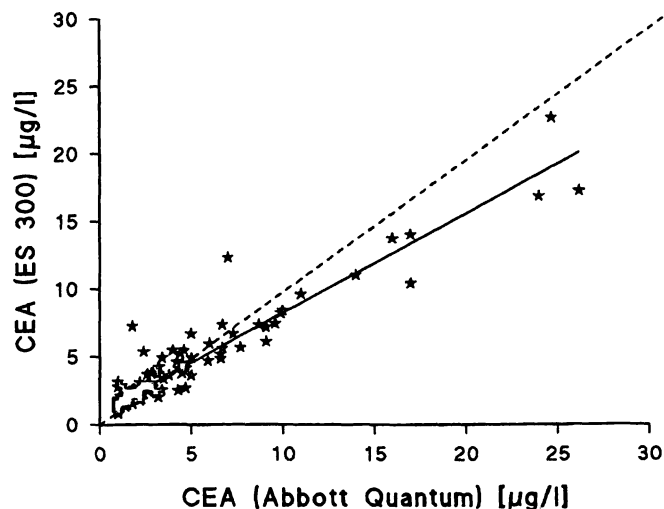


Fig. 1. Correlation between the Abbott immunoassay (x) and ES 300 (y) for CEA. Only values under 30 µg/l ( $n = 131$ ) were used. The line of identity is given as the dotted line, the regression line  $y = 0.48 + 0.83x$  as the solid line, the correlation coefficient  $r$  being 0.98 and the degree of dispersion ( $\tau$ ) 0.74. (This format is used in figures 1–4).

in which CEA was one analyte, up to 4 analytes being measured in one run. Each run included intra- and inter-assay quality assessment samples.

One sample pair showed repeated aberrant results ( $n = 5$ ) with a mean concentration of 7.2 µg/l as measured on the ES 300 and 1.8 µg/l in the Abbott immunoassay. The sample came from a patient operated for colon carcinoma.

Nine samples were assayed as  $< 1$  µg/l by the Abbott Quantum and had values less than 1 µg/l with Enzymun-Test CEA. These were not included in the regression analysis.

Twenty-seven values with an Enzymun-Test System ES 300 concentration  $> 1$  µg/l and an Abbott IEMA concentration  $< 1$  µg/l were included in the regression analysis.

The correlation coefficient of  $r = 0.977$  and  $\tau = 0.74$  was excellent, the regression line  $y = a + bx$  giving values for  $a = 0.48$  and  $b = 0.83$ .

The values in the range  $< 25$  µg/l ( $n = 102$ ) were not significantly different with  $r = 0.945$  and  $\tau = 0.71$ ,  $a = 0.66$  and  $b = 0.79$ .

In the range  $< 15$  µg/l ( $n = 96$ ), the values were identical to those quoted above for  $< 25$  µg/l.

#### Cortisol (fig. 2)

Ninety-six samples were measured in 20 runs (see above). Five samples were discarded due to values lying outside the standard curve.

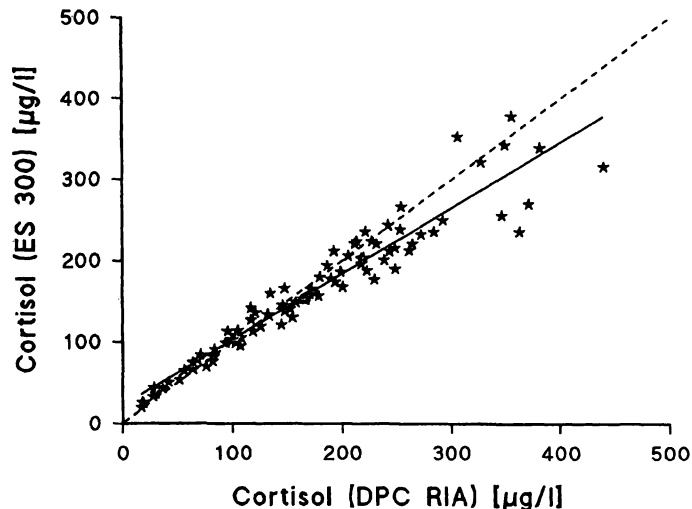


Fig. 2. Correlation between the Diagnostic Products Corporation (DPC) radioimmunoassay (x) and ES 300 (y) for cortisol using 91 data pairs.  $y = 13.2 + 0.85x$ ,  $r = 0.96$ ,  $\tau = 0.89$

Taking all values up to 450 µg/l the regression analysis values were as follows:

Correlation coefficient  $r = 0.96$  and  $\tau = 0.89$ . Slope of the regression line  $b = 0.85$ , intercept  $a = 13.19$ . Again the correlation was excellent.

#### Thyrotropin (fig. 3 – values $< 3.5$ mU/l)

One hundred and fifty-two sera with values under 10 mU/l were measured in 20 assays; 6 data pairs giving results under the detection limit (Enzymun-Test System ES 300  $< 0.07$  mU/l, Behring  $< 0.005$  mU/l) were excluded from the statistical analysis.

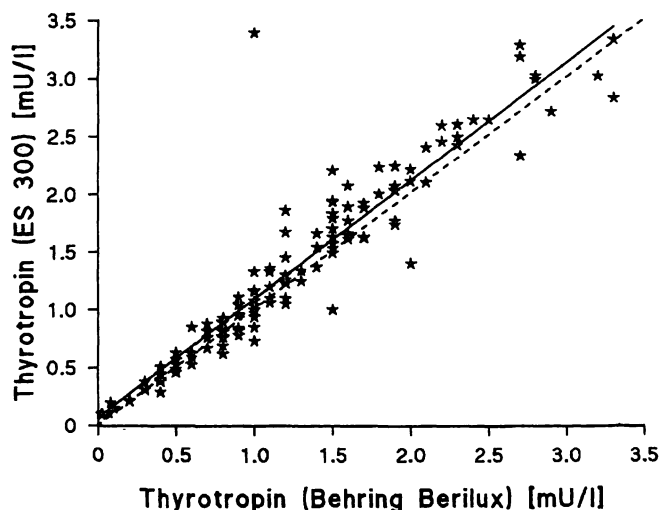


Fig. 3. Correlation between the Behring Berilux immunoluminometric assay (x) and ES-300 (y) for thyrotropin. Only values in the hyper- and euthyroid range ( $n = 146$ ) are shown, as this is the region of interest.  $y = -0.03 + 1.10x$ ,  $r = 0.94$ ,  $\tau = 0.85$

The regression line data for the values up to 10 mU/l were: correlation coefficient  $r = 0.964$ ,  $\tau = 0.85$ , intercept 0.03 mU/l, slope 1.10.

For the 139 values under 3.5 mU/l (i.e. in the euthyroid and hyperthyroid range) the corresponding values were: correlation coefficient  $r = 0.936$ ,  $\tau = 0.85$ , intercept 0.03 mU/l, slope 1.10. The outlier could not be explained (ES 300 = 3.4 mU/l, Berilux 1.0 mU/l) as both  $T_3$  and  $fT_4$  concentrations were within the euthyroid range.

The diagnostic value of both tests was identical, and misclassification errors were not observed.

The lower detection limit of 0.07 mU/l for thyrotropin in Enzymun-Test TSH was perfectly adequate for the differentiation between eu- and hyperthyroid patients.

#### Free thyroxine — $fT_4$ (fig. 4)

One hundred and sixty-one sera were measured. One serum, which was icteric, gave repeatedly different values — with an ES 300 mean concentration 27.4 ng/l, and a Ciba Corning value of 16.3 ng/l ( $n = 5$ ).

Although the correlation between both methods was good ( $r = 0.861$ ,  $\tau = 0.61$ ), the slope of the regression curve showed that the results from the ES 300 were around 40% higher than those measured in the Ciba-Corning LIA, (intercept  $-0.87$ , slope 1.40). These results point to the general problem of standardisation in assays for free thyroid hormones.

The standards for the  $fT_4$  Enzymun-Test were calibrated using values obtained from equilibrium dialysis. Enzymun-Test  $fT_4$  values fitted the clinical pic-

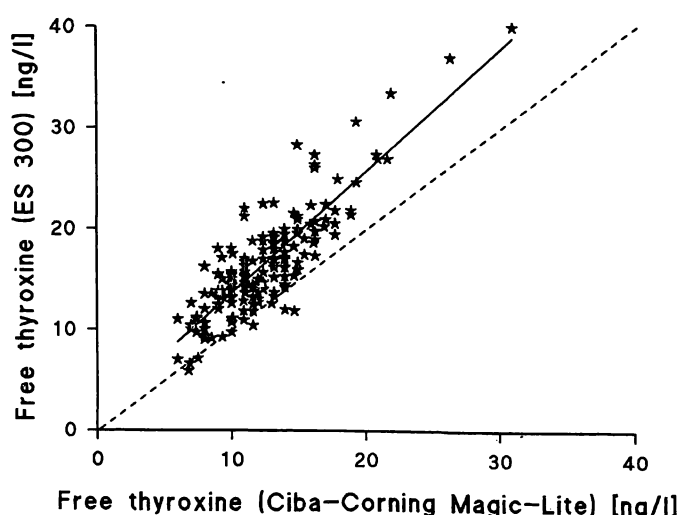


Fig. 4. Correlation between the Ciba-Corning Magic-Lite luminescence immunoassay (x) and ES 300 (y) using free thyroxine values under 40 ng/l ( $n = 161$ ).  $y = -0.87 + 1.40x$ ,  $r = 0.86$ ,  $\tau = 0.61$

ture when they were seen in context with the corresponding Enzymun-Test TSH thyrotropin concentrations.

#### Digoxin

One hundred and forty-four sera were assayed in 20 assays. Four sera with concentrations below the detection limit in both assays were not included in the statistical evaluation. The results were not as good as those for the other analyte comparisons. It is known that the Baxter-Stratus mainly measures non-protein bound digoxin (personal results and l.c. (3)), but this alone cannot account for the discrepancies between ES 300 and Baxter in at least 10 sera, where the values in the Enzymun-Test System method were all higher.

It was not possible to decide which method gave the "correct" values, as a third method (Abbott TDx) agreed with the values from the Baxter-Stratus in 4 cases and with ES 300 in 3 other cases and gave values between both methods in a further 3 cases. The overall correlation was good ( $r = 0.904$ ), even when the dispersion was broader ( $\tau = 0.65$ ), but the intercept (0.15  $\mu\text{g/l}$ ) and the slope (1.06) show a small but constant positive bias in favour of ES 300.

#### Digitoxin

Here 139 sera were measured in 20 assays, 15 data pairs being eliminated as the concentrations measured in both assays lay below the detection limit of the assays. In contrast to digoxin (see above) the correlation between ES 300 and Baxter Stratus was much better, with only 1 discrepant value where a toxic concentration was measured in ES 300 (33  $\mu\text{g/l}$ ) but a therapeutic concentration in the Baxter-Stratus (16  $\mu\text{g/l}$ ). This patient was undergoing chronic renal dialysis.

The overall correlation was good ( $r = 0.943$ ,  $\tau = 0.80$ ) with acceptable intercept (0.08  $\mu\text{g/l}$ ) and slope (0.91).

#### Immunoglobulin E-IgE

From the 133 samples measured, 32 were not included in the statistical evaluation as they had values which were undetectable in both assays.

There were several data pairs which varied widely in both methods, 4 of which differed by more than a factor of 2 at concentrations  $> 200 \text{ kU/l}$ .

The correlation over the whole range was good ( $r = 0.914$ ,  $\tau = 0.77$ ). The in-house method giving values on average 23% lower than the ES 300.

If only the values below 250 kU/l ( $n = 86$ ) are taken into consideration, the correlations is slightly worse ( $r = 0.911$ ), as is the dispersion of the values ( $\tau = 0.73$ ). The values of ES 300 lay on average 16% higher than those from the in-house method. The in-house method was calibrated using an IgE Standard Serum from Behringwerke (Catalogue No. OTFK/02-03). The in-house IgE method was known to give

values consistently lower than the median in an external quality assessment scheme (Wellcome Diagnostics).

### Insulin

The insulin values from 113 out of 115 measured sera were used for statistical purposes. The two data pairs

Tab. 4. Enzymun-Test System ES 300 — Precision data

Analyte	Pool									
	PNIM		PPIM		HS 1		HS 2		HS 3	
	Intra-	Inter-	Intra-	Inter-	Intra-	Inter-	Intra-	Inter-	Intra-	Inter-
<i>Digoxin</i> ( $\mu\text{g/l}$ )										
Mean/Median	1.35	1.31	3.28	3.16	0.70	0.73	1.85	1.84	3.30	3.41
s	0.03	0.05	0.03	0.10	0.02	0.05	0.04	0.07	0.05	0.11
n	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3
CV %	2.2	3.8	3.3	3.3	2.5	7.4	1.9	3.6	1.5	3.2
<i>Digitoxin</i> ( $\mu\text{g/l}$ )										
	Lyphochek 2				Lyphochek 3					
Mean/Median	16.9	17.2	41.6	41.2	6.71	6.59	17.1	16.6	29.6	29.4
s	0.84	0.79	2.33	1.67	0.24	0.28	0.48	0.69	0.85	1.30
n	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3
CV %	5.0	4.6	2.4	4.0	3.6	6.6	2.8	4.2	2.9	4.7
<i>IgE</i> (kU/l)										
Mean/Median	128	121	163	159	63.5	61.2	173	187	326	301
s	3.75	4.53	3.43	6.88	2.06	3.20	6.31	7.57	9.02	11.54
n	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3
CV %	2.9	3.7	2.1	4.3	3.2	5.3	3.7	3.7	2.8	3.8
<i>Insulin</i> (mU/l)										
Mean/Median	20.6	21.1	67.7	73.9	22.0	22.0	47.7	48.1	89.3	93.0
s	1.31	1.56	3.66	5.46	1.09	1.81	1.72	3.30	3.86	4.99
n	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3
CV %	6.4	7.4	5.4	7.4	5.0	8.3	3.6	8.0	4.3	5.4
<i>CEA</i> ( $\mu\text{g/l}$ )										
Mean/Median	4.27	4.26	32.7	32.7	3.82	3.87	16.7	16.6	47.1	47.3
s	0.12	0.18	0.69	1.15	0.16	0.17	0.36	0.45	0.77	1.32
n	2 $\times$ 20	10 $\times$ 3	2 $\times$ 20	10 $\times$ 3	2 $\times$ 20	10 $\times$ 3	2 $\times$ 20	10 $\times$ 3	2 $\times$ 20	10 $\times$ 3
CV %	2.8	4.3	2.1	3.5	3.9	4.5	2.1	2.7	1.6	2.8
<i>Cortisol</i> ( $\mu\text{g/l}$ )										
Mean/Median	125	128	284	284	39.8	41.8	144	144	260	264
s	3.2	4.4	4.5	11.9	2.6	3.5	5.0	10.0	4.5	10.2
n	2 $\times$ 20	10 $\times$ 3	2 $\times$ 20	10 $\times$ 3	2 $\times$ 20	10 $\times$ 3	2 $\times$ 20	10 $\times$ 3	2 $\times$ 20	10 $\times$ 3
CV %	2.5	3.5	1.6	4.2	6.5	8.3	3.5	6.9	1.7	3.9
<i>Thyrotropin</i> (mU/l)										
Mean/Median	3.38	3.39	7.31	7.56	0.53	0.51	2.30	2.24	15.2	15.5
s	0.14	0.13	0.22	0.37	0.02	0.04	0.08	0.09	0.64	0.90
n	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3
CV %	4.0	3.8	3.0	4.9	4.5	6.9	3.6	4.0	4.2	5.8
<i>Free T<sub>4</sub></i> (ng/l)										
Median/Median	10.2	10.4	44.2	43.5	4.10	3.90	14.5	14.1	37.9	37.6
s	0.49	0.73	0.75	0.78	0.35	0.42	0.36	0.59	0.76	0.71
n	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3	20	10 $\times$ 3
CV %	4.8	7.0	1.7	1.8	8.5	10.8	2.5	4.2	2.0	1.9

Intra-assay:  $n = 20$ : the mean value is shown;  $n = 2 \times 20$ : the mean values of both runs are used for calculation.

Inter-assay:  $n =$  number of runs with 3 determination per run; the median values of the three determinations are used for the calculation of the CV.

excluded were under the assay limit in both assays. The results clearly show a standardisation problem in the assays chosen, as the values from the radioimmunoassay from Medgenix lay on average 43% higher than those from the ES-300. The resting insulin values expected in both kits were similar (Enzymun-Test System ES 300 less than 17 mU/l, Medgenix less than 12.5 mU/l) so that the difference in results were not reflected in the reference ranges given by each manufacturer. In taking all 113 data pairs, the correlation was very good ( $r = 0.964$ ,  $\tau = 0.72$ ), the intercept being 3.31 mU/l and the slope only 0.57.

When including only 93 data pairs below 100 mU/l, the picture did not change significantly. The correlation and dispersion coefficients were worse ( $r = 0.909$ ,  $\tau = 0.68$ ) the intercept (3.48 mU/l) and slope (0.57) remaining the same.

#### Precision — Intra- and inter-assay variation (tab. 4 and 5)

The same 5 pool sera were used to establish the intra- and inter-assay variation of the ES 300 in 7 of the assays. In the case of digitoxin, two of the pool sera were changed, as the originals (PPIM, PNIM) did not contain digitoxin as declared analyte. Only values for ES 300 have been taken into consideration. Inter- and intra-assay data for the in-house assays are given for reference purposes.

Table 4 summarizes the results for the 8 analytes. PNIM and PPIM were control sera from Boehringer Mannheim, HS1, HS2 and HS3 were house-internal pool sera, Lyphochek 2 and 3 were from Bio-Rad.

The sera used for the calculation of the intra-assay coefficient of variation were analysed at the beginning and end of the evaluation of the values in the tables  $n = 2 \times 20$  represent the mean of these results.

Tab. 5. Quality control data for the in-house methods used for comparison

Analyte	Range*	Intra-assay CV %**	Inter-assay CV %***
Thyrotropin	0.1 — 50 mU/l	2.9	4.8
Free T <sub>4</sub>	5 — 40 ng/l	3.5	7.1
CEA	3 — 50 µg/l	5.2	8.1
Insulin	10 — 200 mU/l	5.7	8.2
Digoxin	0.6 — 8 µg/l	4.3	5.8
Digitoxin	5 — 90 µg/l	4.2	5.3
IgE	15 — 3000 kU/l	5.1	8.0
Cortisol	10 — 400 µg/l	4.8	6.7

\* Concentration range with mean intra-assay CV under 10%.

\*\* Derived from precision profiles with  $n \geq 200$  data pairs.

\*\*\* Derived from at least 20 assays.

The inter-assay variation was calculated for each of the control sera run at three points in each of 10 runs. The value shown here is the median of the 3 values derived, hence  $n = 10 \times 3$ . Table 5 summarizes the intra- and inter-assay variation of the in-house methods.

The values in the tables show excellent precision within the series as well as between runs, so that values can be determined in singlicate on the ES 300. Even a precision of 11% at a clearly hypothyroid fT<sub>4</sub> concentration of 3.9 ng/l is clinically acceptable, as the result would not lead to a misclassification error as far as diagnosis is concerned.

#### Signal stability over 100 tube runs (fig. 5)

Figure 5 shows the signal drift for thyrotropin and free thyroxine. The same sample was assayed every 4 tubes and the resulting absorbance plotted on the ordinate.

The abscissa shows the tube number. Although inter-sample variations are seen there is no general drift in either of the two assays. Similar results were seen for cortisol and CEA (data not shown).

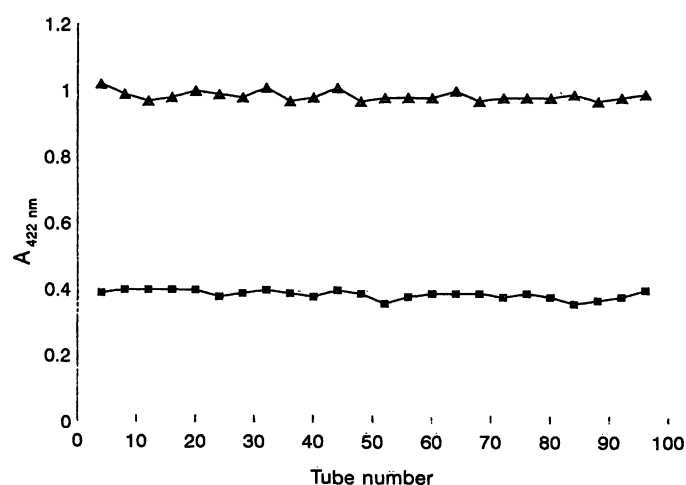


Fig. 5. This figure shows the signal stability over a 100-tube run for thyrotropin ( $\Delta$ ) and free thyroxine ( $\blacksquare$ ). The same sample was placed at 4 tube intervals and was treated as an unknown.

#### Further points

The results presented here show on the one hand the excellent precision expected from a fully-automatic system. On the other hand, they show the variation in results, usually of a systematic (standardisation) nature, encountered in all immunoassays, regardless of analyte or label.

Analytes were chosen to represent immunoassays where "problems" are often encountered, these being

due to cross-reactivity of metabolites, (cortisol, digoxin, digitoxin) (4–6), binding proteins in plasma (fT<sub>4</sub>) (7, 8), lower detection limit (thyrotropin) (9) and immunoglobulin interference (CEA) (10). IgE and insulin were included as analytes which are often measured using radioisotopic labelling.

The precision of the results allows singlicate analysis to be performed, as is to be expected from a fully automated analyser. As stated above, the accuracy cannot be confirmed, because, with the exception of cortisol (11), no reference methods are available for the analytes listed above. Until epitope mapping can be performed and monoclonal antibodies used to define the epitopes which can then be quantified, no progress will be made in this direction, not even with peptide synthesis (12, 13).

The main aim of correlation studies, apart from the correlation coefficient ( $r$ ) and dispersion around the regression line ( $\tau$ ), must be the exclusion of misclassification errors by using the reference ranges set up for the assays used.

The assays tested on the ES-300 showed a good correlation with the in-house methods within the limitations described, and they fulfilled the laboratory criteria with regard to precision and classification of results.

The ES 300, designed as a computer-controlled fully mechanised system, shows distinct advantages when integrated into the laboratory routine.

The interactive dialog on the monitor was seen as positive by laboratory technicians, and the possible use of code words prevents tampering by persons not involved in operation of the ES 300. The computer (EPSON AX-2 with VGA-Monitor (colour) and EPSON LQ-550 Printer, 2 × 5¼" 1.2 MB floppy disks and 40 MB hard disk) is perfectly designed for controlling the ES 300.

In laboratories which have an on-line link to a host computer, all data can be transferred using a bidirectional interface. As this was not the case in this evaluation study, all data were entered by hand. Despite this fact, the ES 300 was fully accepted by the laboratory technicians.

Future developments include the introduction of a bar-code reader which will facilitate communication

at the local level, especially in the entry of patient identity and test-kit characteristics, as well as a software based on Microsoft Windows®. A major advantage of ES 300 was that after pressing the "Run key", the technician could leave the machine.

The computer program allows for run optimisation so that in the case of multi-analyte runs, the incubation/wash steps are so ordered that the total length of time of the run is minimised. Assays which do not meet the internal quality criteria (curve-fit) are not released. The assistant must check all data, and then decide on the course of further actions to reject the results or re-run the test.

The use of a one-point recalibration within a single lot of reagents reduces the costs incurred in manual or semi-automatic methods. The precision allows singlicate determinations and the mean time between service (MTBS) of around 6 months makes the ES 300 an attractive alternative in laboratories performing series of 20–50 patient samples at regular intervals. The "big-brother" Enzymun-Test System ES 600/700 is better suited for longer runs.

The documentation of results can be altered to meet the individual needs of the user so that test reports, sample lists, patient reports and ward lists are all possible.

The use of an enzyme label (peroxidase) and stable chromogen with high molar absorbance (ABTS) (14) reduce the lower detection limit, as well as the time of the colour development step. The use of one enzyme-substrate system also simplifies the assay design as far as automation is concerned. Many reagents are useable in several assays — which cuts down reagent storage space. This tendency will be continued, especially with the introduction of a "universal tube" using the streptavidin-biotin technology. Assays for analytes such as thyrotropin and oestradiol (data not shown here) show that low detection limits are attainable.

The total number of analytes available for the ES 300 is over 30 and covers the fields of endocrinology, tumour markers, viral infections, drug monitoring and other proteins and haptens, so that the majority of analytes in a routine immunoassay laboratory are covered — an important point when considering the introduction of a computer-assisted analyser.

## References

1. Knedel, M., Assmann, G., Courbe, A., van Impe, L., Kattermann, R., Keller, H. & Oellerich, M. (1988) Investigation of the Performance of the ES 600 Enzymun-Test System. A Multicentre Study. *J. Clin. Chem. Clin. Biochem.* 26, 149–162.
2. Passing, H. & Bablok, W. (1983) A New Biometrical Procedure for Testing the Equality of Measurements from two Different Analytical Methods. Application of linear regression procedure for method comparison studies in Clinical Chemistry, Part I. *J. Clin. Chem. Clin. Biochem.* 21, 709–720.
3. Wood, W. G., Färber, P. & Kurowski, V. (1990) A case of divergent digitoxin values under treatment of a patient with acute digitoxin overdose with digitalis antibody fragments. *Klin. Wochenschr.* 68, 324–327.
4. Silber, B., Sheiner, L. B., Powers, J. L., Winter, M. E. & Sadee, W. (1979) Spironolactone-associated digoxin radioimmunoassay interference. *Clin. Chem.* 25, 48–50.
5. Pleasants, R. A., Gadsden, R. H. Sr., Mc Cormack, J. P., Piveral, K. & Sawyer, W. T. (1986) Interference of digoxin like immunoreactive substances with three digoxin immunoassays in patients with various degrees of renal function. *Clin. Pharmacol.* 5, 810–816.
6. Thijssen, J. H., van den Berg, J. H., Adlercreutz, H., Gijzen, A. H. & de Jong F. H. (1980) The determination of cortisol in human plasma: Evaluation and comparison of seven assays. *Clin. Chim. Acta* 100, 39–46.
7. Baldet, L. & Jaffiol, C. (1979) Implications for the measurements of the dialysable fractions of thyroid hormones in thyroid disease. In: *Radioimmunoassay 1979* (Bizollon, C. A., ed.) Elsevier-North-Holland Biomedical Press, Amsterdam, New York, pp. 81–96.
8. Csako, G., Zweig, M. H., Ruddel, M., Glickman, J. & Kestner, J. (1990) Direct and indirect techniques for free thyroxine compared in patients with non thyroidal illness. III Analysis of interference variables by stepwise regression. *Clin. Chem.* 36, 645–650.
9. Mc Bride, J. H., Thiebault, R. V. & Rodgerson, D. O. (1985) Thyrotropin as measured by a sensitive immunoradiometric assay. *Clin. Chem.* 31, 1865–1867.
10. Hansen, H. J., La Fontaine, G., Newman, E. S., Schwartz, M. K., Malkin, A., Möjzisk, K., Martin, E. W. & Goldenberg, D. M. (1989) Solving the problem of antibody interference in commercial "sandwich"-type immunoassays of carcinoembryonic antigen. *Clin. Chem.* 35, 146–151.
12. Siekmann, L. & Breuer, H. (1982) Determination of cortisol in human plasma by isotope dilution mass spectrometry. *J. Clin. Chem. Clin. Biochem.* 20, 883–892.
12. van Leuven, F., Marynen, P., Cassiman, J. J. & van den Berghe, H. (1988) Mapping of structure-function relationships in proteins with a panel of monoclonal antibodies. A study on human alpha 2 macroglobulin. *J. Immunol. Methods* 111, 39–49.
13. Ekins, R. (1991) Immunoassays standardisation. *Scand. J. Clin. Lab. Invest.* 51, Suppl. 205, 33–46.
14. Porstmann, B. & Porstmann, T. (1988) Chromogenic substrates for enzyme immunoassays. In: *Nonisotopic Immunoassay* (Ngo, T. T., ed.) Plenum Press, New York, London, pp. 57–84.

Prof. W. G. Wood  
Institut für Klinische Laboratoriumsdiagnostik  
Klinikum der Hansestadt Stralsund  
Groß Parower Straße 47–53  
O-2300 Stralsund  
Bundesrepublik Deutschland



Eur. J. Clin. Chem. Clin. Biochem.

Vol. 31, 1993, pp. 173–178

© 1993 Walter de Gruyter & Co.  
Berlin · New York

## A Multi-Centre Evaluation of the Measurement of High Density Lipoprotein Cholesterol by the Reflotron® Assay

By *P. Bijster*

*Clinical Laboratory, Martini Ziekenhuis, Groningen, The Netherlands*

### Evaluators

<i>A. P. Anker</i>	<i>Ziekenhuis De Sionsberg, Dokkum</i>
<i>A. J. Bakker</i>	<i>Stichting Klinisch Chemisch Laboratorium, Leeuwarden</i>
<i>P. Bijster</i>	<i>Martini Ziekenhuis (locatie van Swieten), Groningen</i>
<i>J. Bootsma</i>	<i>Refaja Ziekenhuis, Stadskanaal</i>
<i>F. J. Duisterwinkel</i>	<i>Ziekenhuis Ny Smellinghe, Drachten</i>
<i>G. J. Douma</i>	<i>Wilhelmina Ziekenhuis, Assen</i>
<i>P. H. W. van der Ploeg</i>	<i>Martini Ziekenhuis (locatie van Ketwich), Groningen</i>
<i>M. Tel</i>	<i>Streekziekenhuis Oranjeoord, Harlingen</i>
<i>C. G. Thomasson</i>	<i>Ziekenhuis De Tjongerschans, Heerenveen</i>
<i>J. W. E. Visser</i>	<i>St. Antonius Ziekenhuis, Sneek</i>
<i>P. N. Wijnandts</i>	<i>Ziekenhuis Delfzicht, Delfzijl</i>
<i>A. Zwart</i>	<i>Scheper Ziekenhuis, Emmen</i>

(Received July 20/November 23, 1992)

**Summary:** A multi-centre study was undertaken to evaluate the performance of the Reflotron® high density lipoprotein cholesterol method in twelve different clinical chemistry hospital laboratories.

Results were compared with phosphotungstic acid/Mg<sup>2+</sup> and polyethylene glycol/dextran sulphate/Mg<sup>2+</sup> precipitation methods, routinely performed in these laboratories.

Precision was found to be excellent, for both the day-to-day results (SD = 0.03 mmol/l) and for the overall between-laboratory precision (SD = 0.02 mmol/l), measured with control sera.

HDL cholesterol concentrations measured with the Reflotron® system were about 10% lower than those obtained with the two routine precipitation methods, using different instruments with different cholesterol standardization procedures.

### Introduction

After the introduction of the Reflotron® assay for high density lipoprotein cholesterol, which uses dry chemistry reagent and requires no precipitation step (1), reports were published on the evaluation of this method and its comparison with the phosphotungstic acid/Mg<sup>2+</sup> precipitation method (2–4). These evaluations show acceptable linearity (3) and hardly any

interference from drugs or other sources (1, 3). Every reagent strip is magnetically coded with the charge-specific calibration.

During the last decade several comparison studies were made between different precipitation methods (5–8). The major source of error in the separation procedure for HDL in serum is the turbidity of su-

pernatants. Lippi et al. (9) solved this problem for hyperlipidaemic samples by using a mixture of polyethylene glycol, dextran sulphate and  $MgCl_2$  as the precipitation reagent. A test kit with the same reagent concentrations is distributed in the Netherlands and is used in about 40% of the clinical chemistry laboratories in the North of the Netherlands.

It was our aim to compare the results of measurements performed with the Reflotron® analyser with the results obtained with both of the above-mentioned precipitation methods, which are in routine use in twelve different hospital laboratories in the northern region of the Netherlands.

### Materials and Methods

The Reflotron® system (Boehringer Mannheim GmbH, Mannheim, Germany) has been described before and the instruments were operated according to the manufacturer's instructions. EDTA-plasma is used, and precipitation and lipoprotein fractionation by means of dextran sulphate and magnesium ions takes place in the precipitation layer. The precipitation fractions are filtered by the transport layer and only the HDL fraction

reaches the indicator film, where the HDL cholesterol concentration is determined enzymatically by a cholesterol oxidase/peroxidase reaction, measured at 642 nm. The Reflotron® is calibrated against the phosphotungstate method of Boehringer Mannheim.

Twelve Reflotron® systems were installed in twelve different clinical chemical hospital laboratories in the northern region of the Netherlands for a 2-week period (tab. 1). The responsible technicians were instructed by the same Boehringer Mannheim employees. One lot of Reflotron® reagent tabs was used.

Seven of these laboratories (lab 1–7) routinely used a phosphotungstic acid/magnesium ion precipitation method (Boehringer Mannheim No. 543004). Six then used the cholesterol CHOD-PAP method (Boehringer Mannheim No. 642905) on the following mechanized analysers: Hitachi 704 (Boehringer Mannheim), Cobas Fara (Hoffman-La Roche, Basel, Switzerland) and Monarch (Instrumentation Laboratory, Lexington, USA) using different calibration procedures. The other used a CHOD-PAP method manufactured by Technicon (Bayer AG, Leverkusen, Germany, No. 6671) on a RA-1000 (Technicon, Tarrytown, USA). Five of these laboratories (8–12) routinely used a polyethyleneglycol 6000/dextran sulphate/magnesium ion precipitation method (Instruchemie B. V., Hilversum, The Netherlands, No. 2258). Four of these then used the same cholesterol method (Boehringer Mannheim) on the following mechanized analysers: Cobas Mira, Cobas Fara, Monarch and Hitachi 717 using different calibration procedures. The other

Tab. 1. Characteristic features of the phosphotungstate/ $Mg^{2+}$  and the polyethylene glycol/dextran sulphate/ $Mg^{2+}$  precipitation methods used in the 12 hospital laboratories for the routine determination of HDL cholesterol in serum.

Hospital	Routine method	Calibration
1	Phosphotungstic acid/ $Mg^{2+}$ CHOD-PAP (BM) HITACHI 704, 505 nm, 30 °C	Multi-calibrator (BM)
2	Phosphotungstic acid/ $Mg^{2+}$ CHOD-PAP (BM) HITACHI 704, 505 nm, 30 °C	Precimat (BM)
3	Phosphotungstic acid/ $Mg^{2+}$ CHOD-PAP (BM) COBAS FARA, 500 nm, 37 °C	Setpoint (B)
4	Phosphotungstic acid/ $Mg^{2+}$ CHOD-PAP (BM) HITACHI 704, 505 nm, 30 °C	Multi-calibrator (BM)
5	Phosphotungstic acid/ $Mg^{2+}$ CHOD-PAP (BM) COBAS FARA, 500 nm, 30 °C	Multi-calibrator (BM)
6	Phosphotungstic acid/ $Mg^{2+}$ CHOD-PAP (B) RA-1000, 500 nm, 30 °C	Precinorm L (BM)
7	Phosphotungstic acid/ $Mg^{2+}$ CHOD-PAP (BM) MONARCH, 500 nm, 30 °C	Precinorm L (BM)
8	Polyethylene glycol/dextran sulphate/ $Mg^{2+}$ CHOD-PAP (Ic) COBAS MIRA, 500 nm, 30 °C	Artificial cholesterol standard
9	Polyethylene glycol/dextran sulphate/ $Mg^{2+}$ CHOD-PAP (BM) COBAS MIRA, 500 nm, 30 °C	HDL-cholesterol standard (Ic)
10	Polyethylene glycol/dextran sulphate/ $Mg^{2+}$ CHOD-PAP (BM) COBAS FARA, 500 nm, 37 °C	Factor calculated with reference sera
11	Polyethylene glycol/dextran sulphate/ $Mg^{2+}$ CHOD-PAP (BM) MONARCH, 500 nm, 30 °C	HDL-cholesterol standard (Ic)
12	Polyethylene glycol/dextran sulphate/ $Mg^{2+}$ CHOD-PAP (BM) HITACHI 717, 505 nm, 30 °C	HDL-cholesterol standard (Ic)

B = Bayer

BM = Boehringer Mannheim

Ic = Instruchemie

used a CHOD-PAP method from Instruchemie (No. 2142) on a Cobas Mira.

Day-to-day precision of the Reflotron® method was assessed by analysing two control sera, namely Precinorm® HDL-2 with an assigned value of 0.75 mmol/l HDL cholesterol (Boehringer Mannheim) and QCS Normal Unassayed, a human control serum manufactured by Ciba-Corning, Irvine, USA, over a period of 10 days in single mode. Four users of the phosphotungstate/Mg<sup>2+</sup> procedure and four users of the polyethylene glycol/dextran sulphate/Mg<sup>2+</sup> procedure analysed the same control sera with their routine methods over a period of 4–10 days in single mode.

Two different control sera from one lot, namely Precinorm® L with an assigned value of 1.06 mmol/l HDL cholesterol and Precinorm® HDL-1 with an assigned value of 1.09 mmol/l HDL cholesterol, were both distributed to all the laboratories in two bottles with the indications A/C and B/D, respectively. Bottles A and B were analysed in single mode in the first week and bottles C and D were analysed in the second week both on the Reflotron® and with the routine method.

Comparison studies were made by analysing more than 500 fresh patient samples with the Reflotron® and the routine methods in single mode over the period of 2 weeks at the 12 locations. The HDL-cholesterol concentrations covered a concentration range of 0.34 to 2.23 mmol/l.

## Results and Discussion

### Precision and accuracy

The results for day-to-day precision using the two control sera based on human serum (Precinorm® HDL-2 and QCS Nor Unasy) are shown in table 2. The mean within-laboratory coefficients of variation measured at the 12 locations for the two control sera were 4.0% (range 2.3–6.1%) and 2.8% (range 1.5–3.5%), respectively. These results are better than those published by *Leinberger* (1), and the overall mean standard deviation for both these sera (0.03 mmol/l, range 0.01–0.05 mmol/l) is better than the allowable limits proposed by the CDC-NHLBI Lipid Standardization Programme, i.e. 0.08 mmol/l (10).

These results are also better than the mean results measured with the routine methods after phosphotungstate/Mg<sup>2+</sup> and polyethylene glycol/dextran sulphate/Mg<sup>2+</sup> precipitation, which showed coefficients of variation of 4.4% and 4.3%, respectively. Although

Tab. 2. Day-to-day precision results of HDL cholesterol measured in two control sera with the Reflotron® system in the 12 hospital laboratories. Most laboratories also measured the concentrations with their routinely performed precipitation method.

Hospital	Method*	n	Precinorm HDL-2 (assigned value = 0.75 mmol/l)			QCS Normal Unassayed (mmol/l)		
			$\bar{x}$ (mmol/l)	SD (mmol/l)	CV (%)	$\bar{x}$ (mmol/l)	SD (mmol/l)	CV (%)
1	Reflotron	9	0.76	0.04	4.7	1.00	0.03	3.4
2	Reflotron	10	0.75	0.03	4.2	0.97	0.01	1.5
	Phosphotungstate	4	0.87	0.07	8.4	1.04	0.07	6.5
3	Reflotron	10	0.78	0.02	2.3	0.98	0.03	2.6
	Phosphotungstate	10	0.85	0.05	5.7	1.07	0.05	4.8
4	Reflotron	10	0.74	0.03	3.9	0.97	0.03	3.5
5	Reflotron	10	0.77	0.02	3.3	0.94	0.03	3.0
	Phosphotungstate	4	0.84	0.02	2.0	1.02	0.01	1.4
6	Reflotron	10	0.75	0.04	5.5	1.00	0.03	3.5
	Phosphotungstate	10	0.84	0.03	3.8	1.06	0.03	2.7
7	Reflotron	10	0.77	0.05	6.1	—	—	—
$\bar{x}$ Reflotron		69	0.76	0.03	4.3	0.98	0.03	2.9
$\bar{x}$ Phosphotungstate		28	0.85	0.04	5.0	1.05	0.04	3.8
8	Reflotron	10	0.76	0.03	4.6	0.97	0.02	2.3
	Polyethylene glycol	9	0.78	0.01	0.9	1.03	0.02	2.4
9	Reflotron	10	0.78	0.03	3.3	0.99	0.03	2.7
10	Reflotron	10	0.79	0.02	3.2	—	—	—
	Polyethylene glycol	10	0.82	0.05	6.3	—	—	—
11	Reflotron	10	0.76	0.02	2.3	0.97	0.02	2.2
	Polyethylene glycol	6	0.87	0.05	5.9	1.11	0.08	7.1
12	Reflotron	10	0.72	0.03	4.0	0.92	0.03	3.4
	Polyethylene glycol	10	0.85	0.03	3.1	1.06	0.05	4.4
$\bar{x}$ Reflotron		50	0.76	0.03	3.5	0.96	0.02	2.6
$\bar{x}$ Polyethylene glycol		35	0.83	0.04	4.0	1.07	0.05	4.6
Overall $\bar{x}$ Reflotron			0.76	0.02	2.5	0.97	0.03	2.6

\* Phosphotungstate = phosphotungstic acid/Mg<sup>2+</sup>

Polyethylene glycol = polyethylene glycol/dextran sulphate/Mg<sup>2+</sup>

not all the laboratories participated in the routinely performed measurements, and the data for the control sera were obtained over 4–10 days, the results were comparable with the mean coefficient of variation found for the comparison phosphotungstate/ $Mg^{2+}$  method in the Boehringer study, namely 4.4% (1). The mean results of the phosphotungstate/ $Mg^{2+}$  and polyethylene glycol/dextran sulphate/ $Mg^{2+}$  methods for both control sera were 0.08 mmol/l (range 0.07–0.11 mmol/l) higher than the Reflotron® results, which is significant for both methods and both sera ( $p > 0.05$ ).

The mean HDL cholesterol concentration of the Precinorm® HDL-2 measured with the Reflotron® in the 12 laboratories was remarkably close to the Boehringer Mannheim assigned HDL cholesterol value (0.76 mmol/l vs. 0.75 mmol/l), and the overall between-laboratory precision for both control sera was very small, namely 2.5% and 2.6%.

The results obtained with the unknown control sera A–D (Precinorm® L and Precinorm® HDL-1, with Boehringer Mannheim assigned HDL-cholesterol values of 1.06 mmol/l and 1.09 mmol/l, respectively), measured once in both weeks with the Reflotron® and with the phosphotungstate/ $Mg^{2+}$  and polyethylene glycol/dextran sulphate/ $Mg^{2+}$  precipitation methods, can be summarized as follows. The HDL cholesterol overall mean of Precinorm® HDL-1 was almost identical with the Boehringer Mannheim assigned value, namely 1.10 mmol/l vs. 1.09 mmol/l, and was measured with good between-laboratory precision in the 12 laboratories (coefficient of variation = 3.6%). Al-

though the Precinorm® L control serum was not developed for use on the Reflotron® system, and the Reflotron® HDL cholesterol value was therefore far below the Boehringer Mannheim assigned value, this control serum was measured with the same SD = 0.04 mmol/l.

The phosphotungstate/ $Mg^{2+}$  precipitation method, which was used in 7 laboratories, gave a mean HDL cholesterol concentration for the Precinorm® L close to the Boehringer Mannheim assigned value, namely 1.10 mmol/l vs. 1.06 mmol/l, and the precision was excellent (coefficient of variation = 2.1%). These laboratories reported a slightly higher HDL cholesterol concentration for the Precinorm® HDL-1, namely 1.27 mmol/l vs. 1.09 mmol/l, with almost the same excellent precision (coefficient of variation = 1.8%). The mean HDL cholesterol results obtained with the polyethylene glycol/dextran sulphate/ $Mg^{2+}$  precipitation method were almost the same (1.08 mmol/l and 1.25 mmol/l), but the precision was less, with a coefficient of variation of 6.0% and a coefficient of variation of 4.3%.

#### Method comparison

Table 3 shows the comparison of results from the 12 laboratories for the determination of HDL cholesterol by the analysis of EDTA-plasma on the Reflotron® and by the analysis of serum samples by phosphotungstate/ $Mg^{2+}$  and polyethylene glycol/dextran sulphate/ $Mg^{2+}$  precipitation methods. Two hundred and ninety six samples were measured at 7 locations with

Tab. 3. Comparison of results from 12 hospital laboratories for the determination of HDL cholesterol by the analysis of EDTA-plasma samples on the Reflotron® system and by the analysis of serum samples by phosphotungstate/ $Mg^{2+}$  and polyethylene glycol/dextran sulphate/ $Mg^{2+}$  precipitation methods.

Hospital	Method*	n	Range Reflotron (mmol/l)	Slope	Intercept	r	S <sub>yx</sub>
1	Phosphotungstate	35	0.54–1.66	1.23	–0.08	0.932	0.069
2	Phosphotungstate	45	0.53–2.05	1.12	0.00	0.930	0.091
3	Phosphotungstate	52	0.63–2.23	1.16	–0.14	0.939	0.106
4	Phosphotungstate	24	0.65–1.85	1.17	–0.06	0.968	0.064
5	Phosphotungstate	52	0.48–1.68	1.17	–0.06	0.969	0.060
6	Phosphotungstate	51	0.60–2.22	1.05	0.02	0.982	0.045
7	Phosphotungstate	37	0.59–2.12	1.10	–0.01	0.965	0.065
Overall $\bar{x}$ phosphotungstate		296	0.48–2.23	1.14	–0.05	0.955	0.071
8	Polyethylene glycol	48	0.71–1.70	1.26	–0.19	0.970	0.060
9	Polyethylene glycol	47	0.34–1.86	0.98	0.18	0.963	0.054
10	Polyethylene glycol	48	0.66–2.09	1.02	–0.03	0.967	0.068
11	Polyethylene glycol	39	0.59–2.14	1.02	0.01	0.948	0.085
12	Polyethylene glycol	43	0.45–1.46	0.98	0.10	0.965	0.048
Overall $\bar{x}$ polyethylene glycol		225	0.34–2.14	1.05	0.01	0.963	0.063

\* Phosphotungstate = phosphotungstic acid/ $Mg^{2+}$

Polyethylene glycol = polyethylene glycol/dextran sulphate/ $Mg^{2+}$

the Reflotron® system and by phosphotungstate/ $Mg^{2+}$  precipitation, while 225 samples were measured at 5 locations with the Reflotron® system and by polyethylene glycol/dextran sulphate/ $Mg^{2+}$  precipitation.

Regression analysis was done according to the *Passing & Bablok* (11) procedure to determine slope and intercept of the regression line. The Reflotron® vs phosphotungstate/ $Mg^{2+}$  regression analysis give a mean slope of 1.14 (range 1.05–1.23) with a small mean intercept of  $-0.05$  mmol/l (range  $-0.14 - +0.02$  mmol/l) compared with the Reflotron® vs polyethylene glycol/dextran sulphate/ $Mg^{2+}$  results of 1.05 (range 0.98–1.26) and 0.01 mmol/l (range  $-0.19 - +0.18$  mmol/l). These differences are not significant ( $p > 0.05$ ).

It is notable that the laboratory in the group of polyethylene glycol/dextran sulphate/ $Mg^{2+}$  users that applies an artificial cholesterol standard solution for calibration, also reports rather high HDL cholesterol concentrations (above 1.5 mmol/l) compared with the Reflotron® method.

Removing thirteen data points  $> 1.5$  mmol/l, the new regression equation has a slope of 1.16 and an intercept of  $-0.09$ , which results in a total mean slope of the polyethylene glycol/dextran sulphate/ $Mg^{2+}$  group of 1.03 and a mean intercept of 0.03.

The mean slope of the Reflotron® vs phosphotungstate/ $Mg^{2+}$  comparison results in this study is significantly higher than the mean slope found in the Boehringer Mannheim comparison (1) of the Reflotron® and phosphotungstate/ $Mg^{2+}$  method (also performed with 7 laboratories, and with a total of 683 samples), where a mean slope of 1.01 was found. Mean intercept ( $-0.02$  mmol/l) and mean correlation coefficient (0.955) were almost identical.

In other studies, where the Reflotron® HDL cholesterol method (x) was evaluated against Boehringer Mannheim phosphotungstate (y) for EDTA-anticoagulated specimens, somewhat intermediate slopes and comparable intercepts and correlation coefficients were found:

n	Slope	Intercept	r	Reference
100	1.05	$-0.03$	0.977	2
136	1.02	$-0.01$	0.978	3
137	1.08	$-0.12$	0.954	4

To make an overall comparison between both precipitation methods and the Reflotron® method, we pooled the data of 5 laboratories using the phosphotungstate/ $Mg^{2+}$  procedure (3–7) with  $n = 216$  and

range 0.48–2.23 mmol/l, and 5 laboratories using the polyethylene glycol/dextran sulphate/ $Mg^{2+}$  procedure (8–12) with  $n = 225$  and range 0.34–2.14 mmol/l.

The regression line for the Reflotron® (x) vs phosphotungstate/ $Mg^{2+}$  (y) comparison is found to be  $y = 1.12x - 0.03$  ( $r = 0.959$  and  $s_{y,x} = 0.074$ ), and almost the same for the Reflotron® vs polyethylene

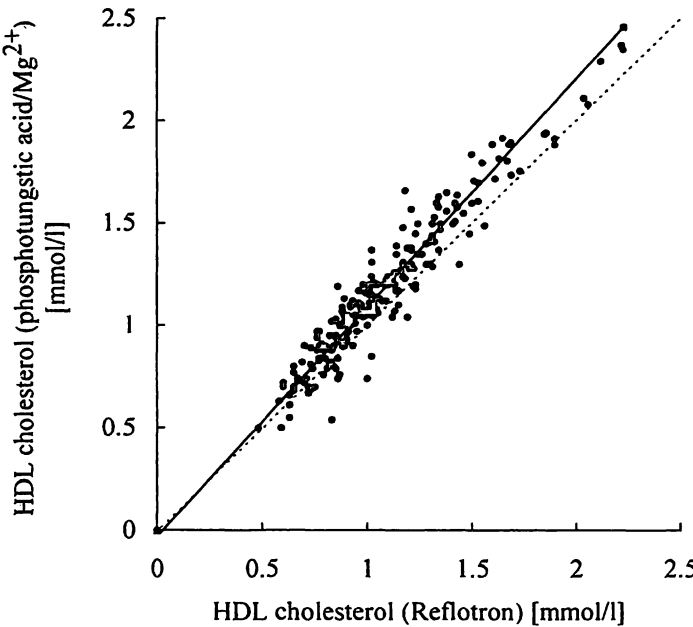


Fig. 1. Comparison of HDL cholesterol concentrations determined with the Reflotron® system (abscissa) and the phosphotungstic acid/ $Mg^{2+}$  precipitation method (ordinate) in 216 fresh EDTA-plasma and serum samples, respectively, in 5 hospital laboratories. Regression equation:  $y = 1.12x - 0.03$ .

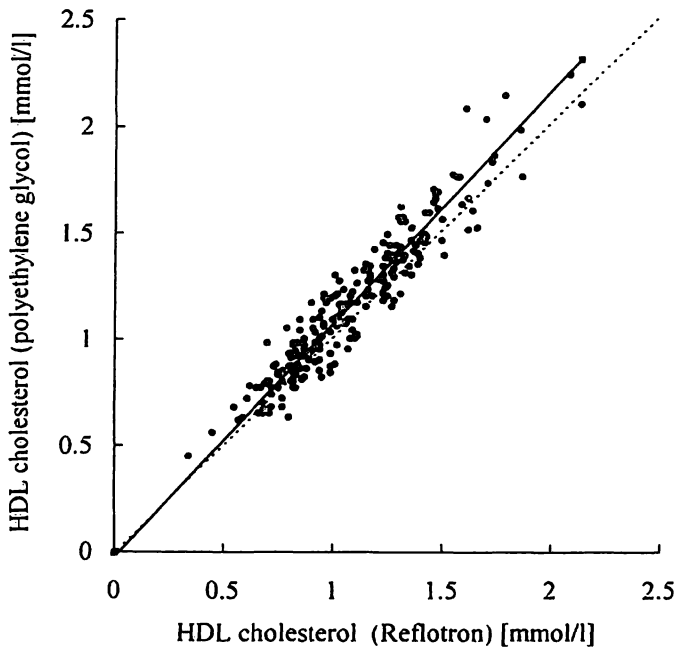


Fig. 2. Comparison of HDL cholesterol concentrations determined with the Reflotron® system (abscissa) and the polyethylene glycol/dextran sulphate/ $Mg^{2+}$  precipitation method (ordinate) in 225 fresh EDTA-plasma and serum samples, respectively, in 5 hospital laboratories. Regression equation:  $y = 1.09x - 0.02$ .

glycol/dextran sulphate/Mg<sup>2+</sup> comparison, namely  $y = 1.09x - 0.03$  ( $r = 0.949$  and  $s_{y \cdot x} = 0.075$ ). These results, also shown in figures 1 and 2, give rise to the conclusion that there is a significant difference between the precipitation methods on the one hand and the Reflotron® method on the other hand, especially for higher HDL cholesterol concentrations.

## Conclusions

From this study, performed in 12 hospital laboratories, it is clear that the HDL cholesterol determination performed on the Reflotron® system with EDTA-plasma samples is a precise and reliable procedure. Overall between-laboratory precision was found to be rather low, compared with the phosphotungstate/Mg<sup>2+</sup> and polyethylene glycol/dextran sulphate/Mg<sup>2+</sup> precipitation methods, routinely performed in these laboratories.

Method comparison showed that HDL cholesterol concentrations determined on Reflotron® were about 10% lower than those reported by the two routine precipitation methods, while these two latter methods gave almost identical regression lines compared with the Reflotron® results. This difference is somewhat higher than found in other comparison studies (1–4). As the routine precipitation methods are somewhat different with respect to instrumentation and chole-

sterol standardization, and this group of laboratories performs well in external schemes (typical between-laboratory standard deviation = 0.04 mmol/l), the overall mean must lie rather close to the "true" HDL cholesterol value for these precipitation methods. This measured difference leads us to question the calibration of the Reflotron® instrument, a subject also discussed in earlier Reflotron® studies (12–13).

Despite these positive results and the fact that the precipitation step has to be performed manually before the sample can be analysed on an automated instrument, the opinion of the participating laboratories is that this Reflotron® method is too time-consuming (90 s per test) and too expensive for implementation in a well equipped clinical hospital laboratory.

This conclusion holds for determinations performed in series, but may be quite different when a rather low sample through-put is analysed, or when a Reflotron® is developed on which more tests can be performed simultaneously.

## Acknowledgement

Thanks are due to Boehringer Mannheim b.v., Almere for providing the Reflotron® instruments and strips.

I also want to thank my colleagues in the Vereniging Klinisch Chemici Noord for participating in this study.

## References

1. Leinberger, R. (1990) Evaluation report Reflotron® HDL cholesterol. Boehringer Mannheim GmbH 3–18.
2. Riesen, W. F. (1990) How reliable is the Reflotron® HDL-cholesterol method? *Schweiz. Med. Wochenschr.* 120, 1971–1975.
3. Ng, R. H., Sparks, K. M. & Statland, B. E. (1991) Direct measurement of high-density lipoprotein cholesterol by the Reflotron assay with no manual precipitation step. *Clin. Chem.* 37, 435–437.
4. Herruer, M. H., Kluitenberg, W. E. & Zuijderhoudt, F. M. J. (1991) Reflotron in lipid screening. Comparison of Reflotron® HDL-cholesterol with a conventional wet chemistry method. *Tijdschr. Ned. Vereniging voor Klinische Chemie* 16, 207–208.
5. Demacker, P. N. M., Vos-Janssen, H. E., Hijmans, A. G. M., van 't Laar, A. & Jansen, A. P. (1980) Measurement of high-density lipoprotein cholesterol in serum: Comparison of six isolation methods combined with enzymic cholesterol analysis. *Clin. Chem.* 26, 1780–1786.
6. Calvert, G. D., Yeates, R. A. & Mannik, T. (1980) A comparison of the phosphotungstate magnesium precipitation and heparin-manganese precipitation procedures for estimating high-density lipoprotein cholesterol. *Ann. Clin. Biochem.* 17, 105–108.
7. Hutt, V., Klör, H. U., Wechsler, J. G. & Ditschuneit, H. (1983) Three different methods for estimating high-density lipoprotein cholesterol. *Clin. Chim. Acta* 128, 173–176.
8. Wiebe, D. A. & Smith, S. J. (1985) Six methods for isolating high-density lipoprotein compared, with use of the reference method for quantifying cholesterol in serum. *Clin. Chem.* 31, 746–750.
9. Lippi, U., Grazianni, M. S., Manzato, F. & Schirella, M. (1986) Assay of cholesterol in high-density lipoprotein in hypertriglyceridemic sera. *Clin. Chem.* 32, 2118–2119.
10. Myers, G. L., Cooper, G. R., Winn, C. L. & Smith, S. J. (1989) The Centers for Disease Control – National Heart, Lung and Blood Institute Lipid Standardization Program: An approach to accurate and precise lipid measurement. *Clin. Lab. Med.* 9, 105–135.
11. Passing, H. & Bablok, W. (1983) A new biometrical procedure for testing the equality of measurements from two different analytical methods. *J. Clin. Chem. Clin. Biochem.* 21, 709–720.
12. Boerma, G. J. M., Gelderland, J., van Gorp, I. & Leynse, B. (1988) Use of the Reflotron® system for cholesterol assay in capillary blood, venous blood and serum – Evaluation of accuracy and lot-to-lot reagent comparability. *Clin. Chem.* 34, 2117–2119.
13. Blijenberg, B. G., Liesting, E. C. & Z'wag, L. (1992) On the accuracy of the Reflotron® creatinine determination. *Eur. J. Clin. Chem. Clin. Biochem.* 30, 51–53.

Dr. P. Bijster  
Klinisch Laboratorium  
Martini Ziekenhuis  
Postbus 30033  
NL-9700 RM Groningen  
The Netherlands

Eur. J. Clin. Chem. Clin. Biochem.  
Vol. 31, 1993, pp. 179–186

© 1993 Walter de Gruyter & Co.  
Berlin · New York

## Evaluation and Performance of the Coulter STKS

By *F. E. A. M. Verheul, J. M. C. Spitters and C. H. F. Bergmans*

*St. Maartens Gasthuis, Venlo, The Netherlands*

(Received March 6/December 21, 1992)

**Summary:** The Coulter STKS is a fully automated haematology analyser which quantitatively measures and computes haematological quantities including a full “five-part” differential. The aim of this study was to evaluate whether the Coulter STKS can be used for screening purposes in a non-academic clinic. Results from the haemocytometry part of the apparatus correlated very well with those from the Coulter S Plus IV. Reasonable to good correlations were obtained with the standard manual 100-cell differential for neutrophils, lymphocytes, monocytes and eosinophils. Significantly higher percentages of monocytes were found by microscopic observation than by the STKS. Samples from 1080 patients were analysed on the STKS, and manual differentials were also performed. After revision of discrepancies and with the exclusion of neonatal samples, a good numerical agreement for the “five-part” differential was obtained. The STKS gave 1.8% false negatives, and both diagnostic sensitivity (96.3%) and specificity (83.3%) were high. A false negative rate of 1.9% was found for the indication leftshift, with a predictive value of a negative test of 97.4%. The predictive value of a positive test, however, was only 40.3%.

An authorization procedure was implemented to reduce false negative samples. The routine laboratory work load for our differentials can be reduced by up to 70%. The Coulter STKS gave a high number of suspect flags for neonatal samples (75%) and is therefore not suitable for screening neonates.

### Introduction

The field of haemocytometry and leukocyte differentiation has changed dramatically in recent years.

Recently equipment has been introduced which is able to perform a five-part differential by means of physical measurements (Sysmex NE 8000, Coulter STKS and the CellDyn 3000CS). The Coulter STKS was introduced in 1989 and combines the Coulter STKR (haemocytometry and three-part differential) with the Coulter VCS (five-part differential). In the VCS part, cells are measured by impedance (V, volume), conductivity (C) and light scatter (S). The computer plots the data two- and pseudo-three-dimensionally. Evaluations of the Coulter VCS showed that the apparatus could be used as a sieve for differentials (1).

This article describes the evaluation of the Coulter STKS in the St. Maartens Hospital at Venlo. Results from the Coulter STKS were compared with those

from the Coulter S Plus IV (haemocytometry) and with the manual differential. Because of practical considerations we chose to compare the STKS differential with the 100 cell microscopic differential. When discrepancies arose, a comparison was made with the 400 cell microscopic differential. A preliminary report has been published elsewhere (2).

We also investigated whether the expectations raised by the evaluation were fulfilled under practical conditions.

### Materials and Methods

#### Instrumentation

The Coulter STKS is a discrete haematology analyser which uses 250 µl EDTA whole blood in the automatic mode, and 150 µl EDTA whole blood in the manual mode. In the automatic mode the apparatus uses up to 12 stackable cassettes with 12 sample positions each.

A small part of all the whole blood (35 µl) goes to the flow-cell, and after lysis of the red cells the white cells are subdivided into 5 populations by 3 different measurements (VCS). The STKS measures about 8000 cells in less than 20 s. The remaining part of the blood is divided between an erythrocyte/thrombocyte and a leukocyte channel. The number of cells is counted by the aperture impedance method. Haemoglobin is measured in a spectrophotometer. The Coulter STKS has a barcode reader to identify samples and it can be linked to a host computer. The capacity of the apparatus for a complete blood determination is 100 samples per hour.

### Sampling

Blood was collected by venipuncture in 2.7 ml Safety-Monovette EDTA tubes (Sarstedt, K<sub>2</sub>EDTA, 1.5 g/l blood). For neonates we used Microtainers with EDTA (Becton Dickinson). When a differential was requested, a wedge film was made immediately after blood collection, and stained with *May-Grünwald-Giemsa* stain in an automated staining machine (Varistain 12, Shandon Elliot).

### Evaluation protocol

As far as possible the protocol of the International Committee for Standardization in Haematology (ICSH) and the U.S. National Committee for Clinical Laboratory Standards (NCCLS) was followed (3, 4). However, because of limited staff numbers, we chose to perform a 100 white blood cell manual differential, and when discrepancies arose between the manual and STKS differential, the manual white blood cell differential was revised by counting 400 cells.

During the evaluation, and later in the daily routine, the software version 1D was used.

### Manual differential

A 100 cell differential was performed as a standard by experienced technologists, who did not have any knowledge of the results of the Coulter S Plus IV or the Coulter STKS. Before the evaluation, criteria were set, according to which a quantitative score was assigned to the respective abnormalities of the manual differential (for details see l.c. (2)). For the abnormalities in the red and white cells a score of 0, 1 or 2 was generally given to the indications 1+, 2+ and 3+. The total score of the manual differential was added and considered as positive if it equalled or exceeded 2. The presence of blasts was always a positive score.

For the indication of leftshift, the following norm was used:

Type of cell	Fraction					
	Sporadic	1–5%	6–10%	11–15%	16–20%	>20%
Band	—	—	1	2	3	4
Metamyelocyte	—	1	2	3		
Myelocyte	1	2	3			
Promyelocyte	2	3				

In addition, the score of toxic granulation and *Döhle* (1+ = 0, 2+ and 3+ = 1) was added. Leftshift was considered as positive if the score equalled or exceeded 2.

### Calibration Coulter STKS

Before the evaluation, the Coulter STKS was calibrated according to the instruction of the company. We did not detect any drift in the equipment.

### Reference values

Blood samples from 206 donors and healthy hospital volunteers in the age range 18–65 years were analysed on the Coulter S Plus IV and the Coulter STKS to calculate the reference values. At the same time a manual differential was performed. Table 1 shows the reference values for the differentials.

Tab. 1. Reference values (95% interval) for differentials

	N	Manual differentiation	Coulter STKS
Neutrophils (%)	191	37.5–73.9	38.7–72.7
Lymphocytes (%)	191	17.2–50.8	21.5–48.3
Monocytes (%)	191	0.7–13.1	2.6–8.2
Eosinophils (%)	194	0–6.5	0–8.2
Basophils (%)	192	0–1.9	0–2.5

### Patient population

During a period of 5 weeks all patient samples with a differential request were analysed. Our goal was to compare at least 1000 patient samples.

### Statistical interpretation of results

Patient data were stored in a HP 9000 Unix computer with the ORACLE relational database management system and calculations were performed with the statistical packets SPSS and Statgraphics.

A 95% reliability interval for the number of counted cells (100 or 400) was assigned to every quantity in the manual differential, according to *Rümke* (5). When the STKS percentage was within this 95% interval the data were considered to be concordant. Orthogonal regression analysis was used for correlation comparisons (6).

## Results

### Analytical aspects, haemocytometry

The analytical aspects were treated in detail in an earlier report (2), in which we showed that samples for differential analysis can be stored up to 8 hours. After 24 hours several suspect flags appeared and division of cell clusters was no longer possible. The between-batch-precision was good. Carry-over was negligible. Linearity was determined with the automated and manual mode. We could not detect any difference between these two modes. A good linearity was found for haemocytometrical quantities. The haemocytometry data of the Coulter STKS correlated very well with those of the Coulter S plus IV, except for the red cell distribution width, for which the STKS gave slightly higher values.

### Comparison of manual differential with Coulter STKS differential;

Table 2 shows the correlation between the percentages of the "five-part" 100-cell manual differential and the



Tab. 2. Correlation of the "five-part" differential microscopic analysis of 100 cells with the results from the Coulter STKS

	N	Regression analysis	r	Reliability interval for mean difference (D)
Neutrophils	197	$y = 1.040 x - 2.450$	0.917	$-2.72 < D < +2.92$
Lymphocytes	200	$y = 1.003 x - 0.135$	0.915	$-2.67 < D < +2.75$
Monocytes	201	$y = 0.667 x + 1.099$	0.550	$-2.75 < D < +2.67$
Eosinophils	198	$y = 0.738 x + 0.922$	0.804	$-0.04 < D < +0.86$
Basophils	191	$y = -0.079 x + 0.768$	-0.031	$+0.25 < D < +0.56$

Random selection from total patient population (with the exception of neonatal samples)

x = microscope differential (100 cell count)

y = Coulter STKS differential

r = correlation coefficient

Tab. 3. Appearance of suspect flags on the Coulter STKS

Suspect flag	Population without neonates	Neonates	Total population
1. Blast	29	11	40
2. Immature granulocytes/bands	139	22	161
3. Variant lymphocytes	16	15	31
4. Review slide	31	11	42
5. Nucleated red blood cells	16	12	28
6. Dimorphous red blood cell population	7	4	11
7. Micro red blood cell fragments	6	5	11
8. Red blood cell agglutination	0	0	0
9. Platelet clumps	2	3	5
10. Giant platelets	2	2	4
Number of patients	1016	64	1080
Patients involved in evaluation	985*	53*	1038*

Number of patients with the exclusion of neonates:

number of neonates:

total patient population:

\* Number of patients (in the respective column) minus suspect flag no. 4 (review slide) in the respective column.

Coulter STKS for a random selection of 200 patient samples. Good correlations were obtained for the neutrophils, lymphocytes and eosinophils. Significantly lower percentages were found on the Coulter STKS for the monocytes. This can be attributed to the unequal distribution of monocytes in the film wedges. The correlation of the basophils was extremely poor, and this may be due to the low number of basophils in the film wedge (reference value to 1.9%).

Table 3 shows the suspect flags which appear in the total population of 1080 patients (several suspect flags are possible per sample). The group of 64 neonatal samples give far more suspect flags (75% of the neonatal samples) than the group of remaining patient samples (almost 20%). Whenever the STKS is used as a differential sieve, the occurrence of a suspect flag must be a warning signal for a manual differential revision.

Agreement in the result of the "five-part" differential was established when the Coulter STKS result was within the *Rümke* interval (5) of the manual differential. For this purpose we first considered the results for the neutrophils, then the lymphocytes, monocytes, eosinophils and finally the basophils. Agreement ought to be present for all 5 cell types. Agreement was found for 779 (79%) of the 985 patient samples (except the neonatal samples). The 206 samples which did not agree were subjected to a 400-cell revision. After revision there was agreement for 903 patient samples (92%). The group of patient samples for which no numerical agreement was found was subdivided. Neutrophils and lymphocytes did not give different results with the two types of differentials. In the manual differential higher percentages of monocytes and eosinophils were found, whereas the Coulter STKS gave higher percentages of basophils.

Agreement normal/abnormal in the patient population (with the exception of neonatal samples)

In order to rationalise differential requests, clear and straightforward indications should be present. In any case, the physician who requests a differential wants to know whether the result of the analysis is normal or abnormal. We have tried to subdivide the results of the differentials in these two groups. A differential result was regarded as normal if the results were within the reference values (see tab. 1), no relevant comment was found in the manual differential (total score 0 or 1), and when the Coulter STKS did not generate suspect flags. The results are described in table 4. After revision of discrepancies only 18 false negatives (1.8%) and 84 false positives (8.5%) were found. The sensitivity and specificity become 96.3% and 83.3% respectively, and the predictive value of a negative result (PV-) on the Coulter STKS becomes high (95.5%). Patients of all ages were present in the group of 18 false negatives. In this group, 16 differentials were judged as being divergent on the basis of

small differences in percentages, and they showed higher percentages or percentages exceeding the reference values in the manual differential (especially for lymphocytes and basophils). For the remaining two patients we found 1.5% myelocytes in the (400 cell) manual differential in one sample (= score 2), while the Coulter STKS did not generate a leftshift suspect flag. In the other patient sample we found 19% blasts in the 100 cell differential (in the 400-cell, 8.8%). The Coulter STKS did not give a blast flag (suspect flag no. 1, tab. 3) or other suspect flag. This patient was known to us with a myelofibrosis.

Table 3 shows that a blast flag was found 29 times in the 1016 patient population. Blasts were found only 3 times in the manual differential (PV+ = 10%). There were 3 false negatives (0.3%), one being the patient with myelofibrosis mentioned earlier, the other two having a suspect flag (immature granulocytes/bands, tab. 3); in table 4, these latter two were included in the column, STKS abnormal. Both these samples had 1% blasts in the 100-cell differential, while no blasts were found in the 400 cell differential.

Tab. 4. Agreement with respect to normal/abnormal in the patient population (with the exclusion of neonatal samples)

		Coulter STKS		
		Abnormal	Normal	Total
Microscope differential (100 cell count)	Abnormal	424	112	536
	Normal	123	326	449
	Total	547	438	985
Sensitivity	79%			
Specificity	73%			
Efficiency	76%			
PV -	74%			
PV +	78%			

The result for the differential is normal if:

- the differential quantities are within the reference values (see tab. 1)
- when there is no relevant comment on the manual differential (total score 0 or 1)
- there are no suspect flags on the Coulter STKS

Tab. 5. Agreement with respect to leftshift in the patient population (with the exception of neonatal samples)

		Coulter STKS		
		Abnormal	Normal	Total
Manual differential (100 cell count)	Abnormal	52	26	78
	Normal	87	820	907
	Total	139	846	985
Sensitivity	67%			
Specificity	90%			
Efficiency	89%			
PV -	97%			
PV +	37%			

Leftshift is present when:

- the score assigned to percentages of the precursors of segments and/or the presence of toxic granulation and *Döhle* is  $\geq 2$ .
- when there is a suspect flag indicating immature granulocytes/bands (tab. 3, suspect flag no. 2) on the Coulter STKS.

One sample had a strong leftshift (score 10!), the other sample only had a raised percentage of segments (93%).

A good indication for a differential is the suspicion of a bacterial infection. For that purpose it is worthwhile to look for a leftshift. Table 5 shows the results of analysis for leftshift (with the exception of neonatal samples). After revision of discrepancies, only 19 false negatives were found (1.9%) with a PV- of 97.8%. The PV+, however, was moderate (40.3%) with a relatively low number of false positives (8.4%). In the group of 19 false negatives there was one patient sample with 11% bands, 2% metamyelocytes and 1% myelocytes and one sample with 13% bands and toxic granulation. It is striking that 12 times over, there was a faulty low percentage of myelocytes ( $\leq 2\%$ ). The other samples only had marginal differences. We do not have data for the other deviations (such as lymphocytosis, abnormal red cells or malignancies) as the number of patients in these groups was too limited.

### Neonatal samples

During the evaluation period a limited number of neonatal samples was analysed. Microtainers containing samples of neonates have to be analysed on the Coulter STKS with the manual mode. Because of the number of samples with a suspect flag "Review slide" (see tab. 3), only 53 neonatal samples were included in a comparison of the white blood cells. The suspect flag "blast" (tab. 3, flag no. 1) was found 11 times, while we could not demonstrate blasts in the corresponding film wedges. We found normoblasts in the 400-cell count, 6 times, whereas the Coulter STKS gave a normoblast warning (erythroblast, tab. 3, flag no. 5) 12 times. For this quantity the PV- was 92% and the PV+ was 17%. The number of false negatives was 6% and the number of false positives 16%.

For the indication leftshift we found 8 positive samples in the 400-cell manual count and 22 positive samples on the Coulter STKS. For the indication leftshift the PV- was 90% and the PV+ was 23%. The fraction of false negatives was 6% and the fraction of false positives 32%. We could not subdivide the neonatal samples into normal/abnormal, because we did not have reference values for this particular population.

### Authorisation procedure

The results of the evaluation were used to draft a flowchart for use in a daily situation. To manage the amount of requests we coupled the Coulter STKS

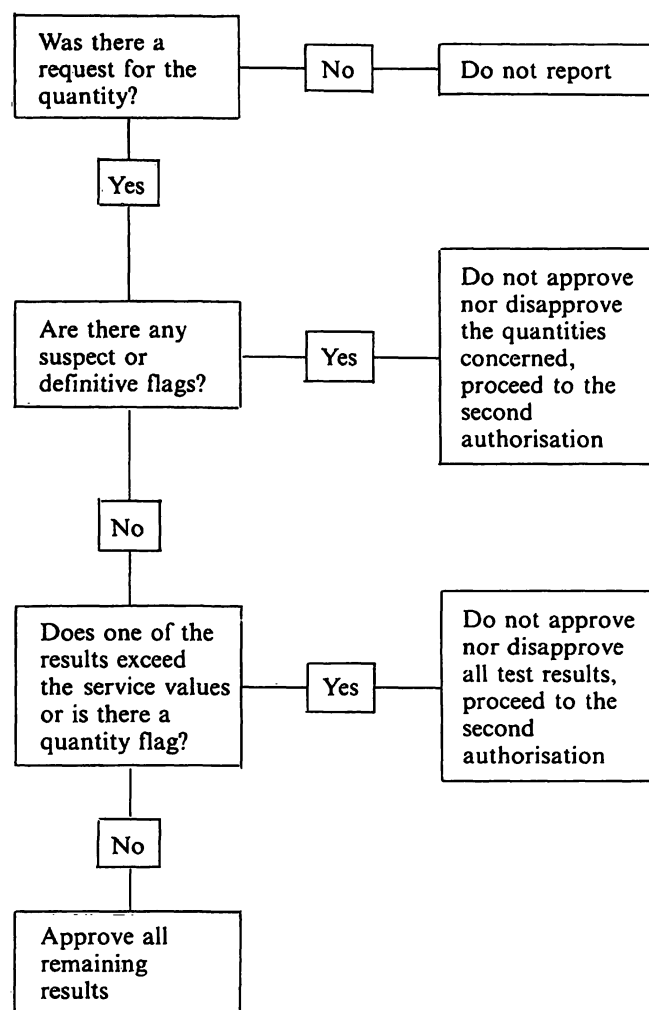


Fig. 1. Flowchart automated authorisation Coulter STKS.

with the host computer via a PC. Samples with barcode labels were placed in random order in the cassettes of the Coulter STKS. Via the PC two types of authorization procedures based on the results of the evaluation were executed. In the first (automatic) procedure, results were compared with fixed criteria. Results of quantities which were not requested were not reported. The fixed criteria are based on definitive and relevant suspect flags and on alert values. Results of samples which did not fulfil the first procedure were evaluated with a second authorization procedure. In this manual procedure the results are compared with earlier data, if present. The flowcharts are shown in figures 1 and 2.

### Performance of the Coulter STKS in actual practice, compared with the evaluation

During a period of 2 months we compared the performance of the Coulter STKS with the results of the evaluation. A total of 4151 patient samples with a differential request was analysed in this period. We expected a reduction of the manual differentials of about 70% for the total population. The number of

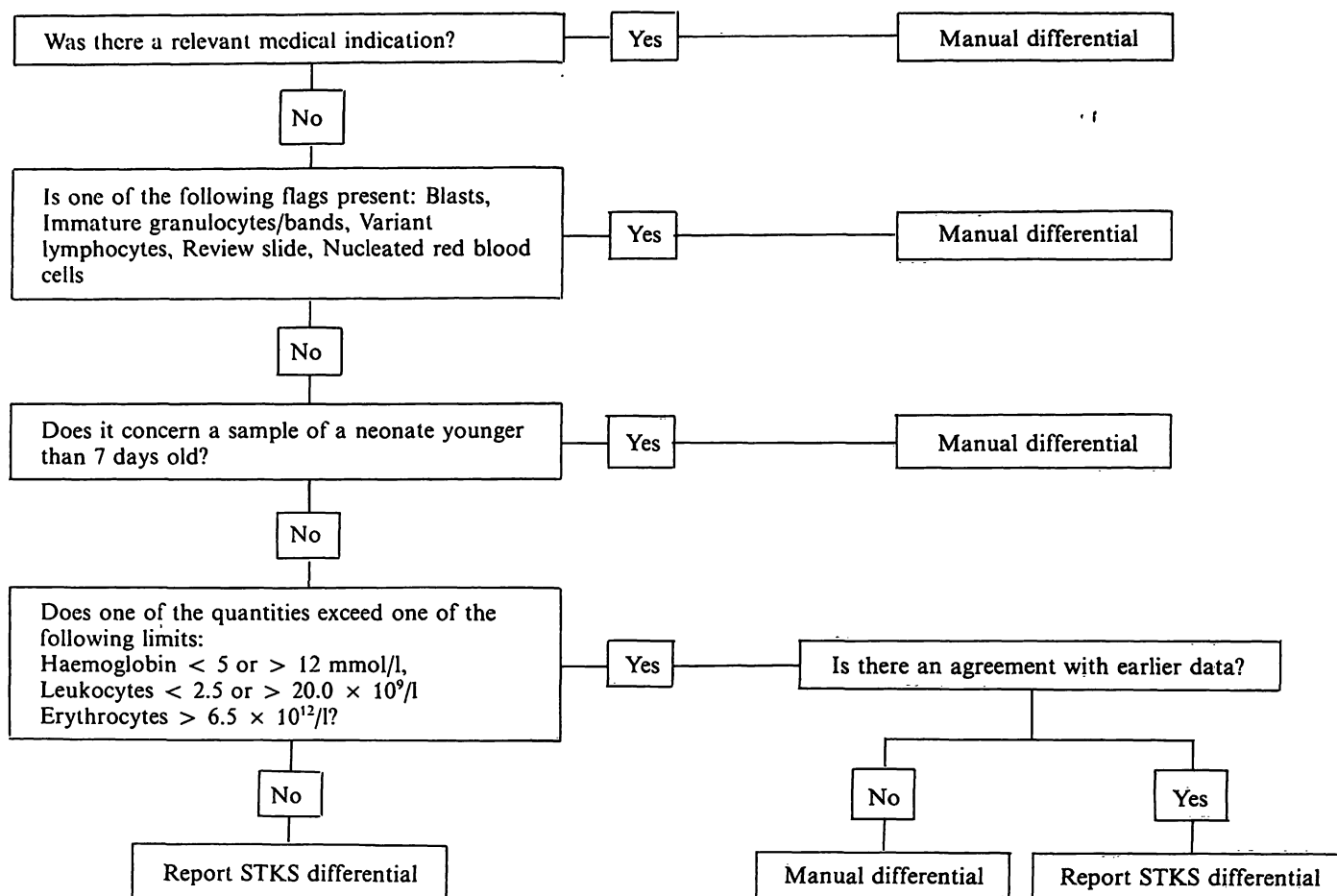


Fig. 2. Flowchart manual authorisation Coulter STKS, decision for manual differential

patient samples with the exclusion of neonatal samples was 3556. From these samples we reported 2738 (77.0%) STKS results and had to perform a manual differential on 818 (23.0%) samples. For 318 (9.5%) reviewed samples we did not find relevant discrepancies, so we reported the STKS results. For the remaining 500 (14.9%) we reported the results of the manual differential.

The number of analysed neonatal samples was 595. We reported the STKS results for 141 (23.7%) samples, and 454 neonatal samples were reviewed. For 122 (20.5%) of the reviewed neonatal samples we did not find relevant discrepancies, and therefore reported the STKS results. The remaining 332 samples (55.8%) were reported according to the manual differentials.

The overall reduction of the manual differentials was 69.4%.

#### What is the relevance of the suspect flags

For the 3556 samples (the neonatal samples were not included) we checked whether the flags "Blasts" and "Variant lymphs" had any significance. The suspect flag "Blasts" was found in 149 samples (4.2%). Upon manual review we detected blasts or normoblasts in

only 4 samples. During the evaluation we found this suspect flag in 2.9% of the samples.

The flag "Variant lymphs" was found in 84 samples (2.4%), and the manual investigation confirmed 73 of these.

#### The Coulter STKS and the problem of the basophils

In a number of cases (about 5%), the Coulter STKS scattergram showed that the subpopulations were not properly separated. In these cases, we found a high percentage of basophils (sometimes even up to 44%). When the samples were re-analysed after 30 minutes, normal percentages of basophils were found. This could be a mixing problem associated with the type of sampling tubes we use (small diameter). If a basophil fraction > 4% was found after re-analysis, the samples were manually reviewed.

#### Discussion

The results of the evaluation indicate that the analytical performance of the Coulter STKS is satisfactory. The stability of the samples for differential purpose

is 8 hours at room temperature. In an evaluation of the Coulter VCS a stability of 5 hours was found (1).

A manual differential based on a 100-cell count gives an analytical variation due to the unequal distribution of cells on the film wedge as well as a variation which can be explained on statistical grounds. In a number of cases the normal physiological variation is greater than changes caused by disease (7). A request for a manual differential is only meaningful when there is a relevant medical indication. The precision of the differential depends on the total number of counted cells (5). Automated differentials have an advantage because they count a larger number of cells (up to 8000).

The STKS results for the neutrophils, lymphocytes and eosinophils correlated well with the results for the manual differential. For the monocytes a fair correlation of 0.550 was obtained. These results are in agreement with other evaluation reports (1, 8). *Barnard et al.* (1) compared the Coulter VCS with the  $4 \times 200$  cells manual differential (4 technologists counting 200 cells each). The Coulter VCS has the option "clinical mode" and "research mode", which enables the user to correct the unequal distribution of the monocytes in the film wedge. The Coulter STKS does not have this option. We found higher percentages of monocytes by manual observation. This is in contrast to the results of *Barnard et al.* (1) for the Coulter VCS, but in agreement with the results of *Warner et al.* (9, 10). The correlation for the basophils was poor. This can be attributed to the low percentage of basophils in the film wedge and also to an incidental higher count of the STKS probably caused by mixing problems.

The numerical agreement of 92% between the STKS and manual differential is good. For that reason the STKS gives good results even for abnormal differentials.

After revision of the discrepancies as shown in the table normal/abnormal (tab. 4) we found 8.5% false negatives and 1.8% false positives. These results are within the recommendations of *Krause* (11) and are comparable with evaluation reports of the Coulter VCS (1), Coulter STKS (10) and Sysmex 5000 (12). *Barnard et al.* found 12.2% false negatives in their VCS evaluation, but they commented that this would have been reduced to 0.9% for a comparison with the

100 cell count. They defined the presence of a deviating cell as abnormal, without considering the number of counted cells. In our study a sample is considered abnormal on the basis of percentages.

The results of our study show that the Coulter STKS can be used for prescreening purposes. Only one false negative sample was missed. To avoid real false negatives we have adapted our procedure, and with the proposal made in the flowcharts we are convinced that there is an extremely low chance of missing abnormal patients.

The STKS is reasonably capable of signalling the presence of a leftshift, but the number of false positives is relatively high. By using the authorization procedure in a daily situation we have reduced the number of unnecessary reviews.

We had a high number of neonatal samples with suspect flags during the evaluation (75%). This is not remarkable, since the software of the STKS Data Management System has been developed for and based on an adult population. The predictive value of a positive test is low, as far as normoblasts and the indication leftshift are concerned, so the number of false positives is high. In a daily situation it means that a large number of neonatal samples has to be reviewed. This was confirmed in the follow-up period.

In our opinion the Coulter STKS is an acceptable haematology analyser which can reduce the daily manual differential work-load by up to 70%. By managing the huge amount of data generated by the Coulter STKS via the flowcharts, and by coupling the STKS with the laboratory host computer, the possibility of false negative results is greatly reduced. We are well aware that a haematology analyser with a full white blood cell differential count can never fully replace the manual differential, and that it will therefore be necessary to maintain relevant and specialised expertise in cell morphology.

### Acknowledgement

The authors wish to thank Mrs. *H. A. W. M. Verhofstad-van de Pasch* for her aid in establishing the criteria for evaluating manual differentials. We would also like to express our gratitude and recognition to all the technologists of the haematology subdepartment who did the additional work during the evaluation with so much devotion. We appreciate the work of Mrs. *J. L. Ebus-Martens* in preparing the manuscript.

## References

1. Barnard, D. F., Barnard, S. A., Carter, A. B., Patterson, K. G., Yardumian, A. & Machin, S. J. (1989) An evaluation of the Coulter VCS differential counter. *Clin. Lab. Haemat.* 11, 255–266.
2. Verheul, F. E. A. M. & Bergmans, C. H. F. (1991) Evaluatie Coulter STKS. *Tijdschr. Nederlandse Vereniging voor Klinische Chemie* 16, 59–67.
3. England, J. M., Rowan, R. M., Assendelft, O. W. van, Coulter, W. H., Groner, W., Jones, A. R., Koepke, J. A., Lewis, S. M., Shinton, N. K. & Thom, R. (1984) ICSH. Protocol for evaluation of automated blood cell counters. *Clin. Lab. Haemat.* 6, 69–84.
4. Kohut, M. L. (1989) An abbreviated method for assessment of electronic differential leukocyte counters. *Labmedica* 6, 27–34.
5. Rümke, C. L. (1983) De nauwkeurigheid van percentages; een nomogram voor de bepaling van de 95%-betrouwbaarheidsintervallen. *Ned. Tijdschr. Geneesk.* 127, 885–888.
6. Cornbleet, P. J. & Gochman, N. (1979) Incorrect least-squares regression coefficients in method-comparison analysis. *Clin. Chem.* 25, 432–438.
7. Costongs, G. M. P. J., Janson, P. C. W., Bas, B. M., Hermans, J., Wersch, J. W. & Brombacher, P. J. (1985) Short-term and long-term intra-individual variations and critical differences of clinical chemical laboratory parameters. *J. Clin. Chem. Clin. Biochem.* 23, 69–76.
8. Jouault, H., Imbert, M., Mary, J. Y., Ade, P., Herpin, J. & Sultan, C. (1990) Différentielle leucocytaire du Coulter VCS. Évaluation et modalités d'utilisation par rapport à l'approche formule Coulter STKR et à la formule manuelle. *Ann. Biol. Clin.* 48, 247–252.
9. Warner, B. A., Reardon, D. M. & Marshall, D. P. (1990) Automated haematology analysers: a four-way comparison. *Med. Lab. Sci.* 47, 285–296.
10. Warner, B. A. & Reardon, D. M. (1991) A field evaluation of the Coulter STKS. *Am. J. Clin. Pathol.* 95, 207–217.
11. Krause, J. R. (1990) Automated differentials in the Haematology Laboratory. *Am. J. Clin. Pathol.* 93, S11–S16.
12. Bürgi, W. & Marti, H. R. (1989) Automated blood count analysis by trimodal size distribution of leukocytes with the Sysmex E-5000. *J. Clin. Chem. Clin. Biochem.* 27, 65–368.

Dr. F. E. A. M. Verheul  
General Clinical Laboratory  
IJsselland Ziekenhuis  
Prins Constantijnweg 2  
NL-2906 ZC Capelle aan den IJssel  
The Netherlands